

AWARD NUMBER: W81XWH-11-1-0550

TITLE: Enhancing the Breadth of Efficacy of Therapeutic Vaccines for Breast Cancer

PRINCIPAL INVESTIGATOR: Jill E. Slansky

CONTRACTING ORGANIZATION: University of Colorado School of Medicine
Auro, CO 80045-2505

REPORT DATE: October 2015

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE October 2015		2. REPORT TYPE Annual		3. DATES COVERED 25Sep2014 - 24Sep2015	
4. TITLE AND SUBTITLE Enhancing the Breadth of Efficacy of Therapeutic Vaccines for Breast Cancer				5a. CONTRACT NUMBER W81XWH-11-1-0550	
				5b. GRANT NUMBER BC100597P2	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Jill E. Slansky, John W. Kappler, Daniel J. Munson, Tullia C. Bruno, Zuly Parra E-Mail: jill.slansky@ucdenver.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Colorado School of Medicine Auro, CO 80045-2505				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The major objective of this project, in collaboration with the other teams at Oregon Health and Science University and City of Hope, is to develop novel strategies to enhance the protective effects of anti-tumor T cells in breast cancer (bc) patients based on the hypothesis that partially protective anti-tumor T cells exist within in most bc patients. We previously determined that the sequence of the TCR of T cell clones expanded in culture must be compared with those isolated ex vivo because their TCR repertoires are skewed in culture. Since the T cells cannot be cultured, we developed an emulsion PCR assay to sequence the alpha and beta chain of single T cells in one reaction. Using this assay we identified TCRs enriched in tumor infiltrating lymphocytes, and found at low levels in the peripheral blood. A panel of TCRs were found in multiple patients suggesting that targeting these TCRs may be most efficient and that they are restricted by the HLA-A2 molecule, the HLA that the bc patients all have in common. With this technique in hand, we are positioned to screen known bc antigens and peptide/MHC libraries for new bc epitopes and mimotopes.					
15. SUBJECT TERMS Breast cancer, tumor antigens, T cell receptor, cancer vaccine					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Unclassified	18. NUMBER OF PAGES 23	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
1. Introduction.....	4
2. Keywords.....	4
3. Overall Project Summary.....	4
4. Key Research Accomplishments.....	18
5. Conclusion.....	18
6. Publications, Abstracts, and Presentations.....	18
7. Inventions, Patents and Licenses.....	19
8. Reportable Outcomes.....	19
9. Other Achievements.....	19
10. References.....	19
11. Appendices.....	NA

1. INTRODUCTION:

We have completed four years of the Multi-Team Project ENHANCING THE BREADTH AND EFFICACY OF THERAPEUTIC VACCINES FOR BREAST CANCER (months 37 to 48). The major objective of this project is to develop novel strategies to enhance the protective effects of anti-tumor T cells in breast cancer patients. This is based on the hypothesis that partially protective anti-tumor T cells exist within most breast cancer (bc) patients. In year 2, we determined that the TCR sequence of T cell clones expanded in culture must be compared and matched with those isolated ex vivo for the identification of epitopes and mimotopes; i.e., the TCR repertoire becomes significantly skewed in culture. In years 2-3, we determined that sequencing the alpha and beta chains separately ex vivo gave no indication as to which chains pairs. So in years 3 and 4, we developed an emulsion PCR assay to sequence the alpha and beta chain of single T cells in one reaction. This method is our most significant reportable out come for years 3 and 4. Using this new method, we identified TCRs enriched in tumor infiltrating lymphocytes and found at lower frequencies in the peripheral blood. A number of TCRs were shared among multiple patients and not control suggesting that these epitopes are HLA-A2-restricted and targeting these TCRs may lead to a more generalizable vaccine than those only good for one patient. These efforts are described in the first half of the report. Using the TCRs identified with this method, we have started to screen a panel of known antigens, peptide libraries, and peptides eluted from breast cancer cells by the OHSU team to determine their cognate antigens. These efforts are described in the second half of the report.

The teams meet regularly on monthly conference call (organized by the OHSU group) and in person bi-annually. Since our last report, we met on 2-4-2015 at the City of Hope and on 9-1-2015 in Denver.

2. KEYWORDS

Breast cancer, tumor antigens, T cell receptor, cancer vaccine

3. OVERALL PROJECT SUMMARY

Task 1. Generate reagents and identify conditions for experiments to follow: months 1- 40, Lee, Slansky, and Spellman This project summary is arranged in the order of the Statement of Work (revised 1-1-12). Due to the nature of the project, we are still performing task 1, details are included in the other tasks. The tasks that we have not included are either being performed by the other team, or have not been started yet.

Task 2. Enroll 100 patients with all major breast cancer subtypes from City of Hope and University of Colorado.

The following table is a list of samples that we obtained over the last 4 years from Dr. Virginia Borges at the UC Anschutz Medical Campus (AMC, [Part 1, Table 1](#)) and Dr. Peter Lee at the City of Hope ([Part 2, Table 1](#)).

Table 1. Patient samples obtained from AMC (Part 1) and City of Hope (Part 2)

Patient*	Subtype	Stage	HLA type	CD8 yield	CD4 yield	Sample use of CD8 TILs**
UC0147	Basal-like	Stage I T1N0M0	HLA- A*02:01	117,750	92,250	Used to analyze function of TIL and troubleshoot sequencing
UC0152	Luminal A	Stage I T1N0M0	Not done	Not done	Not done	Tissue too small, not used
UC0157	Basal-like	Stage II T2N0M0	HLA- A*02:01	<6,000	<6,000	Used to troubleshoot sequencing
UC0197	Luminal B	Stage I T1N0M0	HLA- A*02:01	<4,000	<4,000	Sequencing
UC0198	Luminal B	Stage II T2N1M0	HLA-A*23	68,000	24,000	TIL RNA stored for future use
UC0200	Luminal A	Stage III T3N3M0	HLA- A*02:01	T:30,000 LN: 37,500	T:15,000 LN: 62,500	Sequencing, TDLN included
UC0202	Luminal A	Stage II T2N1M0	HLA- A*01:01	<4,000	<4,000	TIL RNA stored for future use
UC0205	Luminal A	Stage II T2N1M0	HLA- A*11:02	<6,000	<6,000	TIL RNA stored for future use
UC0211	Luminal A	Stage I T1N0M0	HLA- A*02:01	16,000	Not done	Sequencing
UC0213	Luminal A	Stage II T2N0M0	HLA- A*24:02	Not done	Not done	Tissue had too much lipid, no selection
UC0217	Basal-like	Stage I T1N0M0	HLA- A*29:02	12,000	6,000	TIL RNA stored for future use
UC0221	Basal-like	Stage II T2N0M0	HLA- A*03:01	100,000	94,000	TIL RNA stored for future use
UC0238	Luminal B	DCIS only	HLA- A*01:01	Not done	Not done	Used to trouble shoot IFN-g capture assay
UC0255	Unknown	Stage I T1N0M0	HLA- A*02:01	Not known	Not known	Used to troubleshoot sequencing
UC0256	Basal-like	Stage I T1N0M0	HLA- A*02:01	256,000	76,000	Sequencing
UC0277	Basal-like	Stage I T1N0M0	HLA- A*02:01	Not done	Not done	Used to troubleshoot sorting specific populations for single cell PCR
UC0308	Luminal A	Stage II T1N1M0	HLA- A*02:01	8,000	4,000	Emulsion PCR, single cell sequencing
UC0323	Luminal B	Stage II T2N0M0	HLA- A*02:01	60,000	24,000	Emulsion PCR, single cell sequencing
UC0336	Luminal A	Stage I T1N0M0	HLA- A*02:01	24,000	12,000	Emulsion PCR, single cell sequencing

*Other patients were consented, but we did not obtain tissues

**RNA from CD4 TILs is being stored for future use

Patient	Subtype*	Stage*	HLA type	CD8 received	CD4 received	Sample use of CD8 TILs and LN**
C	Luminal	Stage	HLA-	T: 10,000	T: 10,000	Sequencing comparison of

	ER/PR+	IIB	<u>HLA-A*02:01</u>	LN: 10,000	LN: 10,000	<i>ex vivo</i> and <i>in vitro</i> expanded TIL
D	Luminal ER/PR+	Stage IIB	<u>HLA-A*02:01</u>	T: 10,000 LN: 10,000	T: 10,000 LN: 10,000	Sequencing comparison of <i>ex vivo</i> and <i>in vitro</i> expanded TIL
E	Luminal ER/PR+	Stage IIA	<u>HLA-A*02:01</u>	T: 10,000 LN: 10,000	T: 10,000 LN: 10,000	Sequencing comparison of <i>ex vivo</i> and <i>in vitro</i> expanded TIL
BC41	Luminal A	Stage IIIA	<u>HLA-A*02:01</u>	T: 30,000 LN: 30,000 PBL: 220,000	T: 30,000 LN: 30,000 PBL: 250,000	Emulsion PCR, single cell sequencing
BC48	Luminal ER/PR+		<u>HLA-A*02:01</u>	LN: 10,000 PBL: 20,000	LN: 30,000 PBL: 30,000	Emulsion PCR, single cell sequencing
BC35	Luminal B	Stage IIA	<u>HLA-A*02:01</u>	T: 30,000 PBL: 80,000	T: 14,000 PBL: 80,000	Emulsion PCR, single cell sequencing
BC47	Luminal ER/PR+		<u>HLA-A*02:01</u>	LN: 30,000 PBL: 30,000	LN: 30,000 PBL: 30,000	Emulsion PCR, single cell sequencing
BC61	Luminal A	Stage IIA	<u>HLA-A*02:02</u>	T: 40,000 LN: 40,000 PBL: 40,000	T: 40,000 LN: 40,000 PBL: 40,000	Emulsion PCR, single cell sequencing
BC57	Luminal A	Stage IIB	<u>HLA-A*02:01</u>	T: 16,500 LN: 10,000 PBL: 30,000	T: 12,000 LN: 30,000 PBL: 30,000	Emulsion PCR, single cell sequencing
BC55	Her2+	Stage IV	<u>HLA-A*02:01</u>	T: 30,000 PBL: 30,000	T: 30,000 PBL: 90,000	Emulsion PCR, single cell sequencing
BC54	Luminal A	Stage IIIA	<u>HLA-A*02:01</u>	T: 10,000 PBL: 20,000	PBL: 20,000	Emulsion PCR, single cell sequencing
BC67	Luminal A	Stage IIIA	<u>HLA-A*02:01</u>	T: 20,400 LN: 20,600 PBL: 100,000	T: 23,000 LN: 37,000 PBL: 100,000	Emulsion PCR, single cell sequencing
BC70	Basal	Stage IV	<u>HLA-A*02:01</u>	T: 50,000 PBL: 50,000	T: 30,000 PBL: 20,000	Emulsion PCR, single cell sequencing
BC71	Basal	Stage IA	<u>HLA-A*02:01</u>	T: 30,000 LN: 50,000 PBL: 50,000	T: 30,000 LN: 100,000 PBL: 100,000	Emulsion PCR, single cell sequencing
BC80	Luminal ER+PR+	Stage IIA	<u>HLA-A*02:06</u>	T: 5,000 LN: 25,000 PBL: 25,000	T: 5,000 LN: 25,000 PBL: 25,000	Emulsion PCR, single cell sequencing
BC81	Luminal ER+PR+	Stage IIA	<u>HLA-A*02:01</u>	LN: 25,000 PBL: 25,000	LN: 25,000 PBL: 25,000	Emulsion PCR, single cell sequencing
BC85	Luminal ER+	Stage IIA	<u>HLA-A*02:01</u>	T: 10,000 LN: 10,000 PBL: 50,000	T: 13,000 LN: 17,000 PBL: 50,000	Emulsion PCR, single cell sequencing
BC75	Luminal ER+PR+	Stage I	<u>HLA-A*02:01</u>	T: 5,000 LN: 100,000 NT: 3,000 PBL: 100,000	T: 5,000 LN: 100,000 NT: 5,000 PBL: 100,000	Emulsion PCR, single cell sequencing
BC86	Luminal ER+PR+	Stage IA	<u>HLA-A*02:01</u>	T: 47,500 LN: 50,000	T: 50,000 LN: 50,000	Emulsion PCR, single cell sequencing

				PBL: 50,000	PBL: 50,000	
BC87	Luminal ER+PR+	Stage IIIA	HLA- A*02:06	T: 50,000 LN: 50,000 PBL: 50,000	T: 50,000 LN: 50,000 PBL: 50,000	Emulsion PCR, single cell sequencing
BC92	Luminal ER+PR+	Stage IA	HLA- A*02:06	T-L: 10,000 T-R: 22,000 LN-L: 60,000 LN-R: 100,000 PBL: 50,000	T-L: 7,000 T-R: 14,000 LN-L: 100,000 LN-R: 100,000 PBL: 50,000	Emulsion PCR, single cell sequencing
BC94	Luminal ER+PR+	Stage IA	HLA- A*02:01	LN: 30,000 PBL: 50,000	LN: 50,000 PBL: 50,000	Emulsion PCR, single cell sequencing
BC96	Luminal ER+PR+	Stage IIA	HLA- A*02:01	T: 10,000 PBL: 100,000 NT: 4,000	T: 20,000 PBL: 100,000 NT: 6,000	Emulsion PCR, single cell sequencing
BC124	HER2	Stage IIIB	HLA- A*02:01	LN: 500,000 PBL: 500,000	LN: 500,000 PBL: 500,000	Emulsion PCR, single cell sequencing
BC113	Luminal ER+PR+	Stage IIA	HLA- A*02:01	T: 50,000 PBL: 100,000	T: 90,000 PBL: 100,000	Emulsion PCR, single cell sequencing
BC132	Luminal ER+PR+	Stage IIIA	HLA- A*02:01	LN: 12,000 PBL: 95,000	PBL: 31,000 PBL: 400,000	Emulsion PCR, single cell sequencing

Task 3e: Process patient samples (blood, TDLNs, tumor) and expand T cells ex vivo [to obtain the their TCR sequence].

Although expansion of T cells in culture provides us with more cells for analysis, the population dynamics changes such that same highly frequent clones present in the pre-expansion population are lost in the post-expansion population. Additionally, some V regions as well as VJ combinations dominate the post-expansion population that are negligible in the pre-expansion population. We determined in year 2 that while between 29-63% of clones differ by <10 fold between the populations, there is no guarantee that these are tumor specific clones; even though some clones are highly frequent in the pre-expansion population and remain so following expansion. Furthermore, it is well established that tumor resident T cells are often dysfunctional, and when removed from the tumor microenvironment do not proliferate and instead die. Thus, we pursued direct ex vivo sequencing. We started these studies sequencing the alpha and beta chains separately summarized last year. At the time of this report these data are with Peter Lee's group awaiting publication. Identification of the alpha and beta chains together better addressed our ultimate goals of identifying epitopes and mimotopes for bc immunotherapies. The following is a draft of the manuscript describing this method and our results using emulsion PCR of bc samples and controls.

Identification of shared TCR sequences from human breast cancer

Introduction

Infiltration of breast tumors by alpha-beta CD8+ T cells is associated with better outcomes and longer survival times for breast cancer patients (1,2). Targeted immune-based therapies hold great promise toward improving breast cancer treatments. Numerous studies have examined the immunophenotype of breast cancer infiltrating T cells (TILs), suggesting that activated non-suppressive T cells are of most benefit. Further assessment of the tumor infiltrating T cell (TIL) repertoire has broad implications

for breast cancer therapies in antigen discovery, cancer vaccines, and adoptive cell therapies.

Unique genetic recombination events produce both the alpha and beta chains of the T cell receptor (TCR). These chains heterodimerize to generate the diversity of the T cell repertoire and antigen specificity of each TCR. Many studies have analyzed the alpha and beta chains of TIL TCRs separately [reviewed in (3)]. High throughput sequencing of individual chains has begun to describe the diversity of TIL (4). However, these data do not harbor the alpha-beta pairing information. Pairs are readily identified after expansion of T cell clones, although culture of T cells leads to substantial skewing of the repertoire (5), which may select for T cells of varied affinity or avidity. Single-cell sequencing identifies alpha-beta pairs, but is often laborious and has relatively low throughput. An exciting high throughput statistical method to determine the matched alpha-beta pairs using bioinformatics was published recently, which is optimal for hundreds of thousands of TCRs (6).

Methods using emulsion PCR increase the sensitivity and fidelity of standard PCR reactions, and were recently applied to T cell repertoire analysis of bulk peripheral blood CD8 T cells (7). This technique facilitates the analysis of large numbers of TCR pairs from single cells. However, the technique as published is limited to a single V-beta region and a limited number of V-alpha regions. To elucidate the T cell repertoire responding to breast cancer, we required a method that included all 93 potential TCR V genes.

We developed an expanded emulsion RT-PCR protocol so we could sequence the CD8+ T cell repertoires of tumors and peripheral blood (PBL) from HLA-A2+ breast cancer patients, and PBL from control donors. Using this method on hybridomas with known TCRs, we showed that 85% of the alpha and beta chains paired with the expected partner and hybridomas that made up 1% of a pool were readily identified. Using this method with primary tumor or blood T cells of unknown identity or diversity, we identified on average 10^3 unique sequences per 10^4 T cells. Finally, we identified multiple shared alpha-beta pairs among patient tumors that were not present in control samples. Of these, some were exclusively identified within the tumor repertoire suggesting tumor-directed clonal expansion and providing potential tumor-specific TCRs.

Results

To identify TCR sequences of TIL, we modified the emulsion RT-PCR protocol of Turchaninova et al (7) to include primers for all V-alpha and V-beta TCR gene fragments. These primers included overlapping complementary regions to facilitate annealing of the appropriate alpha and beta gene fragments to each other during PCR. Following purification of DNA from the emulsion, a PCR reaction using nested C-region primers and blocking oligos with non-complementary 3' ends was used to amplify alpha-beta joined products and suppress amplification of unpaired gene fragments, respectively. A second nested PCR was performed to add the adaptors and barcodes for Illumina sequencing. For analysis of the sequence reads of the paired alpha-beta TCRs, we used a modified version of MiTCR (8) which we termed CompleteTCR.

To determine the robustness of this protocol, we mixed 10 different hybridomas expressing a variety of alpha and beta genes prior to running them in the emulsion RT-PCR protocol. Replicate samples were processed without the emulsion using the same cycling parameters for all steps and sequenced in order to determine pairing fidelity and efficiency (Figure 1). The frequency of TCRs that were paired with the correct input TCR was approximately 85% when the emulsion was used, and that number dropped to

between 5 and 10% when no emulsion was present. Analysis of the incorrect pairs showed no one alpha chain was dominantly paired with any given beta chain or vice versa, suggesting random pairing in the absence of the emulsion (data not shown).

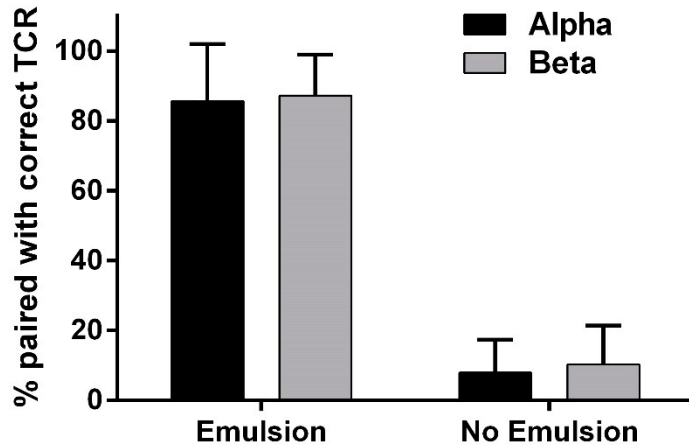


Figure 1. TCR pairs are recovered with the correct match at a high frequency in the emulsion RT-PCR. Ten 5KC T cell hybridomas were mixed and emulsion RT-PCR was performed either with or without emulsion phase. Sequences were then analyzed for pairing with the correct alpha or beta TCR. Percentages represent the number of reads which paired with the input alpha or beta out of the total number of reads for the analyzed alpha or beta. $n = 3$ replicate experiments. Error bars represent standard deviation from the mean. The average read count was 670,000 over the three experiments.

To confirm that all 92 V-alpha and V-beta primers amplified a corresponding TCR fragment, we performed real-time PCR reactions using these primers and pooled cDNA derived from the peripheral blood T cells of 4 donors. Each of the reactions yielded a product within 15-35 cycles, showing that this primer set binds and amplifies all potential alpha and beta TCR genes.

We next applied this technology to identify shared T cell clonotypes in tumors of breast cancer patients. Patients were consented, HLA-typed, and their tissue samples were collected (Table 2). CD8+ T cells were positively selected from digested tumors and peripheral blood, incubated overnight in IL-2 to increase the concentration of intracellular TCR transcript, and subjected to emulsion RT-PCR. On average we identified 10^3 unique sequences per 10^4 cells (Figure 2A). The pairing fidelity of the most frequently identified TCR pair is over 90% for the beta chain and 87% for the alpha chain (Figure 2B). The fidelity of beta to alpha pairing was lower than alpha to beta since often an excluded non-functional alpha TCR chain is transcribed but not expressed, whereas only one beta transcript is transcribed.

Patient ID	Age	Stage	TMR Grade	Sentinal LN status	ER	PR	Her2	HLA-A alleles
p1	59	IIIA	III	+	+	+	-	02:01
p2	49	IIA	III	ND	+	+	+	02:01, 03:01
p3	55	IIB	II	-	+	+	-	01:01, 02:01
p4	54	IIIA	II	+	+	+	-	02:01
p5	66	IIIA	II	+	+	+	-	02:01, 01:01
p6	44	IA	II	+	+	+	-	02:01, 29:01
p7	30	IA	III	-	+	+	+	02:01, 24:03
p8	35	IA	III	-	+	+	-	02:01, 31:01
p9	45	IIA	II	+	+	+	-	02:01, 03:01
p10	53	IA	II	-	+	+	-	02:01, 03:01
p11	47	IA	II	-	+	+	-	02:01, 03:01
p12	53	IIA	III	-	+	+	-	01:01, 02:01
p13	56	IV	II	ND	-	-	+	02:01, 03:01
p14	71	IIA	III	-	+	-	-	02:01, 03:01
p15	63	IV	II	ND	-	-	-	02:01, 32:02
p16	60	IA	III	-	-	-	-	02:01
p17	57	IIA	III	-	+	+	-	02:02, 30:02
p18	29	IIA	III	-	+	+	-	02:06, 33:01
p19	42	IIIA	III	+	+	+	-	02:06, 74:01
p20a	53	IA	I	ND	+	+	-	02:06, 74:01
p20b			II	ND	+	+	-	

Table 1. Patient demographic information for samples collected at The City of Hope Beckman Research Institute and the University of Colorado. ND = not determined

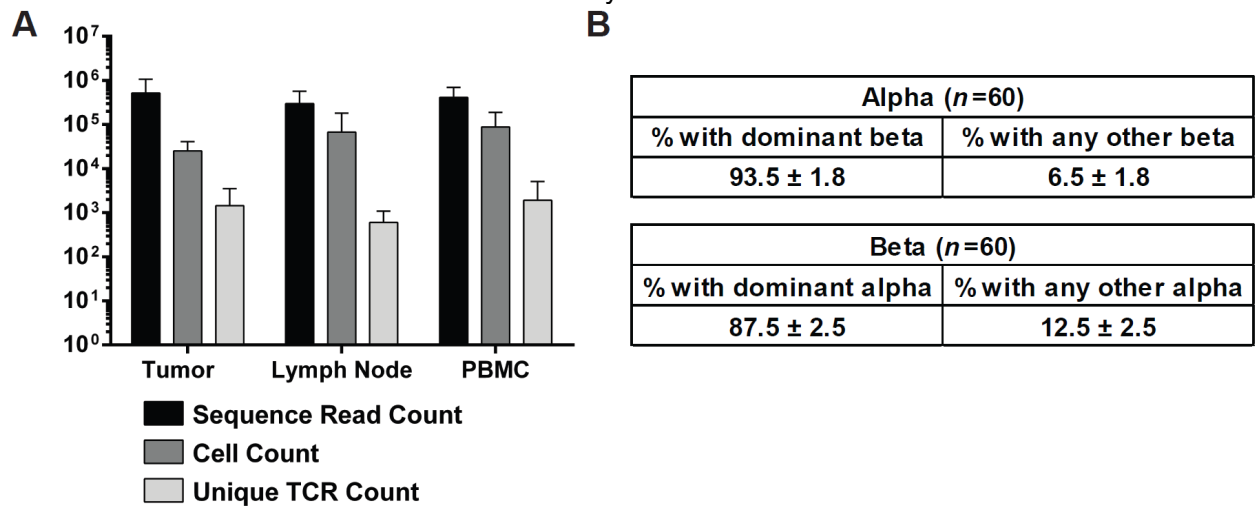


Figure 2. Pairing efficiency of tumor derived TCRs exhibits predominant pairing with a single partner. Sequence read counts, cell counts, and unique TCR pairs are shown for each tissue component analyzed (A). Highly frequently represented TCRs were analyzed for the percentage of partnership with a particular pair and the frequency with which it paired with any other TCR (B).

Using the emulsion RT-PCR protocol, we determined the heterogeneity of the T cell repertoire in tissues from distinct spatial regions of a patient who presented with one tumor in each breast (p20, [Table 1](#)). We collected a peripheral blood sample and tissue from the tumors and sentinel lymph nodes from each side. The alpha-beta pairs identified from the tumor, LN, and peripheral blood were ranked on a yellow to black scale corresponding to their recovered sequence frequencies ([Figure 3A and B](#)). Since the breast cancer subtype and stage were similar between the samples, we expected the TIL TCRs from each side to also have similar sequences. Of the samples sequenced, 29 shared TCR pairs were identified between the tumor repertoires out of 280 (left breast tumor) and 250 (right breast tumor) unique TCR pairs. A larger overlap may be detected with larger sample sizes of the blood or LN; however, these results show significant overlap in TCR sequences from one patient suggesting reproducibility in the method when analyzing primary TIL. Additionally, 92% of TCRs present in the left tumor and 83% in the right tumor were absent in the respective LN and blood ([Figure 3C](#)). Although we expected a larger degree of overlap with the sentinel LN and the tumor repertoires, most of the shared TCRs were not found in the blood.

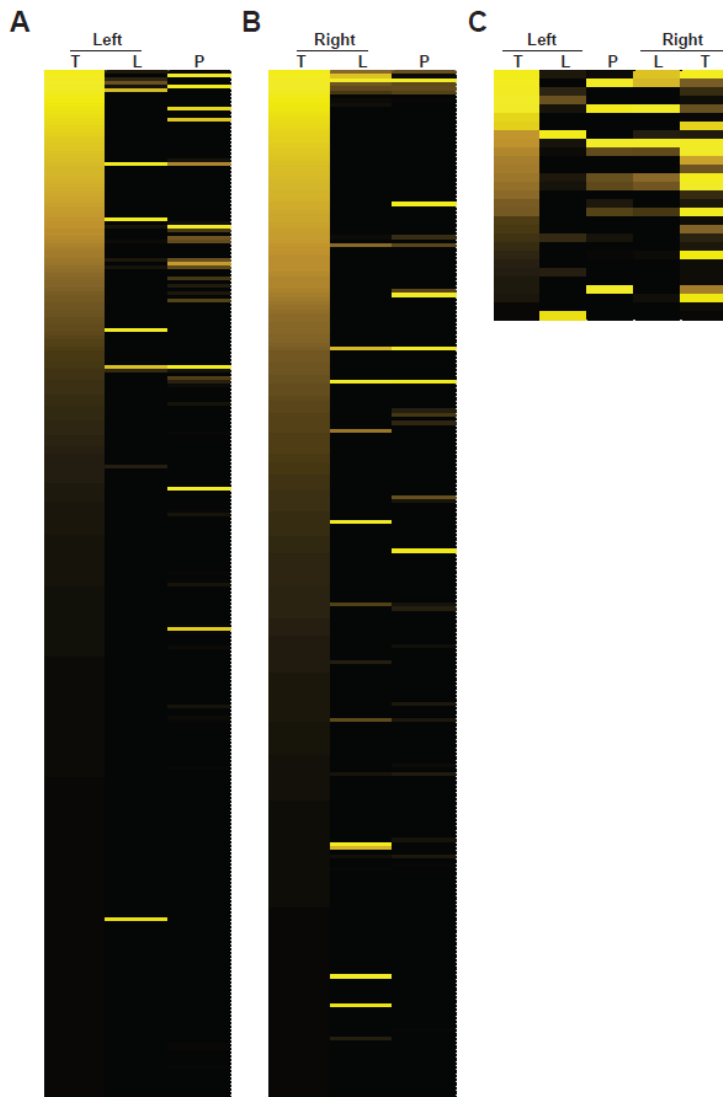


Figure 3. Comparison of tumor derived TCRs from the same patient reveals overlap in the responding TCR repertoire. BC patient #16 presented with tumors in each breast. The tumors were resected along with the sentinel LN from each axilla. Rank heat map of all tumor derived TCRs from the left (A) and right (B) tumor and the corresponding ranks in the sentinel LN and blood. C. Ranks of the tumor TCRs shared between each tumor and corresponding ranks in the LN and blood, and isolated to show direct comparison between the tumors.

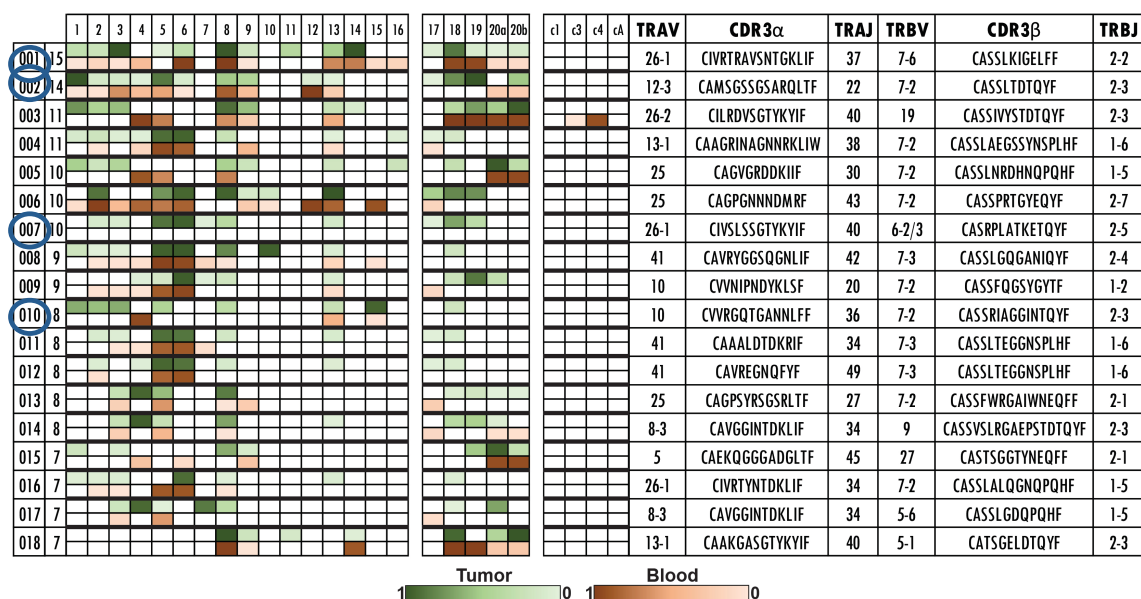


Figure 4. Sharing of TCR pairs across breast cancer tumors reveals a public response among patients with a shared HLA type. TCR pairs that were shared between 7 or more patients are listed (right columns) with the number of patients (left column) where the specific TCR pairs was identified. For each pair, a color value corresponding to the rank value of the TCR in that sample is shown for the tumor (green) and blood (red).

With the goal of ultimately activating a breast cancer-specific immune response in many recipients, we next determined which TCRs from the tumor samples were shared across patients, but not the tumor-free controls. We hypothesized that tumor-specific T cells are more frequently identified in TIL than blood, so we also compared the sequences to those from the PBL. Shared TCRs from 20 patient tumors and corresponding peripheral bloods were determined (Figure 4). The TCR pairs TRAV26-1:CIVSLSSGTYKYIF:TRAJ40;TRBV6-2/3:CASRPLATKETQYF:TRBJ2-5 and TRAV10:CVVNIPNDYKLSF:TRAJ;TRBV7-2:CASSFQGSYGYTF:TRBJ1-2 (blue circles, Figure 4) were identified in the tumor, but not in the corresponding blood samples or the controls. These are tantalizing candidates for the development of tumor-specific vaccines. There will be additional analysis here regarding the false discovery rates and statistical likelihoods that the TCRs enriched in the tumor samples are in fact tumor specific. These analyses are currently being done at OHSU with their experience in bioinformatic comparisons.

Discussion

To identify TCR pairs from breast cancer patients as eventual targets of immunotherapy, we modified and expanded a previously published method of single-cell emulsion RT-PCR (7). With this new method, we analyzed the T cell repertoire ex vivo, unaffected by skewing or cell death during culture. Single-cell emulsion RT-PCR provides information about the identity of individual T cells by analyzing the TCR chains as a pair, eliminating the need for computer-based algorithmic pairing or mix-and-match of separate alpha and beta chains. This high-throughput technique can be applied to large diverse populations, limited numbers of cells, or low diversity populations.

We verified the pairing efficiency of emulsion RT-PCR using murine 5KC hybridomas transfected with human TCRs. Pairing efficiency, the frequency one chain is identified with the expected chain, is approximately 85% in this system. However, emulsion RT-PCR lacks a quantitative aspect, presumably due to primer binding differences, transcript abundance and/or stability, or cDNA conversions. The pairing efficiency from T cells from primary tissue is approximately 90%, although the calculation to obtain this number requires the assumption that a given alpha or beta only pairs with one other chain in the repertoire. Thus, this estimate is likely low. Because the composition of TCR chains in vivo is unknown, how quantitative this analysis is also unknown, and the rank values should be viewed as an approximation. To precisely determine the population dynamics of a given emulsion RT-PCR-derived pair, this technique could be paired with more established TCR profiling methods such as single-cell RT-PCR from sorted cells, cloning of tumor T cells, and/or RNA or DNA-based bulk sequencing.

Analysis of two spatially distinct tumors from the same individual revealed 10% homology in the T cell repertoires. Because of tumor intrinsic factors contributing to the efficiency of obtaining sequences of the alpha-beta pairs such as conditions of the tumor (hypoxia, necrosis, fat, other factors), cell viability, and sample size we were surprised that the amount of overlap was relatively high. Although further analysis of the TCR target of the homologous TCRs is needed, these results suggest that a shared response to tumor does exist.

We obtained around half a million sequences per sample; 1-3% of the TCR sequences appeared once and were dropped from the analysis because we were ultimately interested in the T cells that have expanded in response to the tumor. Dropping the sequences also avoids analysis of sequencing errors. Although the average number of T cells with the same TCR is unknown, we obtained an average of 1 unique TCR per 10 T cells in the reaction. Published reports have suggested the beta chain TCR to T cell ratio to be 1/1040 (9). Additionally, it has been shown that the size of the TCR beta repertoire in an individual is $\sim 3 \times 10^6$ drawing from an estimated 5×10^{11} possible sequences (10). We recovered a lower diversity than has been described potentially due to inefficiencies during the RT-PCR process such as low transcript abundance, primer biases, and reverse transcription efficiency.

The analysis of the alpha-beta pairs from breast cancer TILs and blood shows a predominance of TRBV7-family member TCRs, suggesting that this V-beta family is responsible for tumor recognition or the primers confer bias due to mispriming, as there are 8 primers for the TRVB7 family in these assays (Figure 4). However, there are also 8 primers for the TRVB6 family and we only identified one shared TCR utilizing this V-beta gene. Use of V region primers designed to bind the upstream conserved region and sequencing technology that delivers longer read lengths may become a more attractive option to identify closely related family members than multiple primers.

We discovered alpha-beta pairs shared among 7-15 of 20 patients, but not control blood from female HLA-A2+ or HLA-A2- donors without breast cancer (Figure 4). Most of the shared TCRs were frequently identified in the tumor repertoire(s) but absent or infrequent in the matching blood repertoire, suggesting that these TCRs may recognize the tumor. These may be tumor-specific T cells in the periphery, non-tumor-specific memory T cells responding to the inflammatory environment of the tumor, or T cells in the tumor vasculature.

Most of the patients (16 of 20) and all controls had at least one copy of HLA-A*02:01, the most common allotype in Northern Asian and North American populations. The other 4

samples also expressed alleles of the HLA-A*02 supertype, suggesting that these alleles bind to the same peptides (11). Thus, it is likely that the shared TCR is restricted by the shared HLA allele. Since 6 of 20 patients also shared HLA-A*03:01+, we will also consider its restriction for the relevant TCRs. Future efforts will determine the HLA-restriction and cognate peptide antigens of these TCRs. The patient samples collected to date were identified by their HLA-A2 serotype. Understanding the antigens responsible for the expansion of these T cells may contribute to future antigen-specific therapies.

Materials and Methods

Cell lines

All 5KC T cell hybridomas were propagated in S-MEM supplemented with essential and non-essential amino acids, Pen/Strep, 7 mM NaHCO₃, 75 µM gentamycin, 750 µM sodium pyruvate, 3 µM dextrose, 1.7µM L-glutamine, 36 nM BME, and 10% FCS. The 5KC cell lines used are listed in Supplementary Table 1 and were a kind gift from Dr. Andrew Fontenot.

Tissue Procurement and Processing

Fresh tumor and lymph node tissues were obtained from patients who gave Institutional Review Board-approved consent prior to tissue collection at City of Hope Beckman Research Institute and the University of Colorado Health. Tumor tissue and tumor-infiltrated lymph nodes were separated from fat tissue and minced into pieces up to 2 mm in diameter with scalpel blades. Fragments were then treated with 0.2 Wunsch U/ml Liberase (Roche) and 10 units/ml DNase (Sigma) in RPMI (Life Technologies) for up to 1 h as needed until the tissue dissociated. Enzymatic dissociation was stopped by adding 5 ml of RPMI and 10% FBS. The digested tissue suspension was then filtered through a 70-µm filter followed by a 40 µm filter. Red blood cell (RBC) lysis was performed if necessary using RBC Lysis Buffer (Biolegend). Lymph nodes without tumor infiltration were similarly treated, but without enzymatic digestion. Peripheral blood samples were Ficoll separated (GE Healthcare) to remove cellular debris and RBCs. T cell populations were isolated from single-cell suspensions using a CD8 Positive Selection Kit (Stem Cell Technologies) following manufacturers recommendations and incubated in 10 IU/ml IL-2 (Peprotech) overnight prior to emulsion RT-PCR.

HLA typing

We collected peripheral blood from patients at the time of consent, 1-2 weeks prior to surgery. After Ficoll separation, we determined the HLA serotype by flow cytometry with an HLA-A2 antibody (clone BB7.2, Biolegend). We isolated genomic DNA by TRIzol extraction from 1x10⁶ mononuclear cells for high-resolution genotyping from the HLA-A2+ samples (HLA Lab, City of Hope).

Emulsion RT-PCR

An emulsion phase was generated using the Micellula Emulsion and Purification Kit, per the manufacturer's recommendations (EURx), adapted from Turchaninova et al (REF). 300 µl of emulsion phase was added to 50 µl of RT-PCR master mix [(1x Qiagen OneStep RT-PCR buffer, 1x Qiagen Q-Solution, 0.4 µM dNTPs, 1µl Qiagen OneStep RT-PCR enzyme mix, 10 U of RNaseOUT (Life Technologies), 600 nM C region primers, 600 nM Stepout primer, 60 nM V region primers (Supplemental Table 2)] containing 5,000-250,000 T cells and vortexed on high for 5 min at 4°C. The resulting 350 µl reaction was separated into 3 PCR tubes and subjected to the following RT-PCR

regimen: 1 cycle of 65°C for 2 min; 1 cycle of 50°C for 40 min; 1 cycle of 95°C for 15 min; 40 cycles of 94°C for 30 sec, 60°C for 1 min, 72°C for 2 min; and 1 cycle of 72°C for 5 min. The 3 PCR reactions per original emulsion phase were then pooled and purified using the Micellula Emulsion and Purification Kit.

Nested PCR

Two nested PCR reactions were performed following RT-PCR. The first nested PCR amplified the entire eluted RT-PCR product in 100 µl total volume using Taq polymerase (New England Biolabs), 0.2 µM primers (Supplemental Table 1) and the following cycling conditions: 95°C for 30 sec; 30 cycles of 95°C for 30 sec, 52°C for 30 sec, 68°C for 1 min; and 1 cycle of 68°C for 5 min. The second nested PCR was performed using LA Taq (New England Biolabs) and 2 µl of the first nested PCR and 0.4 µM primers (Supplemental Table 1) in 50 µl total volume with the following cycling conditions: 94°C for 30 sec; 20 cycles of 94°C for 30 sec, 53°C for 30 sec, 65°C for 1 min; and 1 cycle of 65°C for 5 min. The resulting products were gel purified and quantified using a QuBit Fluorometer (Life Technologies).

Illumina High Throughput Sequencing

Separate samples were pooled in equal molar amounts and adjusted to a final concentration of 1 ng/µl. 6 µl of the pooled sample was diluted 1:1 with 0.2 M NaOH and incubated at room temperature for 5 min prior to addition of 1 ml of Hybridization Buffer (Illumina). A 6:10 dilution of the sample was then made in Hybridization Buffer and 600 µl of the diluted sample loaded on an Illumina MiSeq flow cell and paired end 2x250 sequencing was performed per the manufacturer's instructions.

Bioinformatic processing

Raw sequence reads were processed using the open-source Galaxy Platform (refs below). Reads were paired by joining the 3' end to create a 500 nt read with the small portion of the C regions on either end. The reads were then clipped from either end using the engineered overlap sequence as the identifier. This created two separate files, each containing only TCR information from either the alpha or beta TCR. Non-clipped reads were discarded. The clipped files which contained identical cluster IDs for paired alpha and beta TCRs were then analyzed for their TCR identities using CompleteTCR. The CompleteTCR pipeline was built from MiTCR (8), an efficient tool for CDR3 extraction, clonotype assembly, and repertoire diversity estimation, but limited to independent analysis of the alpha or beta TCR chain. CompleteTCR allows determination of alpha-beta pairs by manipulating raw MiTCR outputs using an R script¹. Two modest changes to the MiTCR source code were made, since MiTCR assigns each input read a numeric identifier. First, the standard MiTCR results file was written to include a list of the numeric IDs for all reads belonging to each pair. Second, a temporary output file was created to map the sequence identifier for each read in the input FASTQ file to its MiTCR-assigned numeric identifier. No changes were made to the algorithms MiTCR uses for CDR3 extraction, clonotype assembly, or error correction. The R script first annotated the reads of each alpha clonotype with the appropriate sequence identifiers, repeating the process for the reads of each beta clonotypes. The alpha and beta reads were next paired by their sequence identifier, and any read lacking a mate was removed from the dataset. Finally, the frequencies of alpha-beta pairs were

¹ R Core Team (2014): A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available: <http://www.R-project.org/>.

calculated. CompleteTCR required Java version 1.7.0² or higher and R version 3.1.0¹ or higher with the plyr package version 1.8.3³ or higher. It was run from the command line via a shell wrapper script that required an input manifest detailing locations of the alpha and beta FASTQ files and their corresponding sample names. These results were further analyzed using standard procedures in Excel.

We performed the following analysis to determine rank of each TCR sequence in a sample:

$$\text{Rank} = \frac{(P^t + 1) - P^n}{P^t}$$

Where P^t is the total number of unique frequency positions and P^n is the position number of interest. One was added to the total position number to make the range of ranks from 0-1. Position was determined by assigning an increasing value to decreasing unique frequencies, such that CDR3 sequences with identical frequencies would be assigned identical positional identifiers.

Task 4. Identify and isolate anti-tumor T cells from TDLNs and tumor samples: months 1-40, Lee, Slansky, and Spellman

4b. Generate specific tumor antigen lysates from recombinant baculovirus-infected insect cells. *Currently expressing antigens in C1R cells since these cells can also be used as antigen presenting cells.

We will screen T cells against a panel of known BC antigens prior to screening peptide libraries. If the T cells react with these known antigens, the road to antigen and mimotope discovery will be greatly shortened. Toward this goal, we have stably transfected C1R:A2 cells (from Vic Engelhard at University of Virginia) with known breast cancer antigens HER2, NY-ESO-1, hTERT, CEACAM5, and MUC-1. Evidence of tumor antigen expression is suggested by detection of Thy1.1 and breast cancer antigen-specific antibody staining. Although these cells lines were re-derived in the last year, the data are similar to those in last year's report.

To confirm that the specific HLA-A*02:01-restricted tumor antigens are processed and presented by these cells, we are generating antigen-specific TCR transfectomas. We are sub-cloning the control TCRs and those from Figure 4 into hybridoma cell lines to make transfectomas, and use these to indicate antigen recognition on the C1R:A2-tumor antigen cell lines. This methodology will continue to be helpful when screening the peptide library.

The generation of the TCR transfectomas is detailed in **Figure 5**. Briefly, the TCR sequences are purchased and inserted upstream of the mouse constant regions. The product is transduced into a T cell hybridoma (5KC) that has little to no natural TCR expression, but expresses human CD8. The plasmids also contain GFP used for cell

² Java SE - Downloads | Oracle Technology Network | Oracle (n.d.). Available: <http://www.oracle.com/technetwork/java/javase/downloads/index.html>. Accessed 6 May 2014.

³ Wickham, Hadley (2011) The Split-Apply-Combine Strategy of Data Analysis. J Stat Softw 40: 1–29.

sorting in combination with the surface expression of the TCR molecule. TCR functionality can be assessed by binding to a potential target and measuring the release of IL-2 by ELISA. **Figure 6** shows that we have successfully generated transfectomas for TCRs 002 and 010 from **Figure 4** and a control TCR that recognizes NY-ESO-1 bound to HLA-A2. The summary of our progress in making transfectomas is shown in **Table 3**.

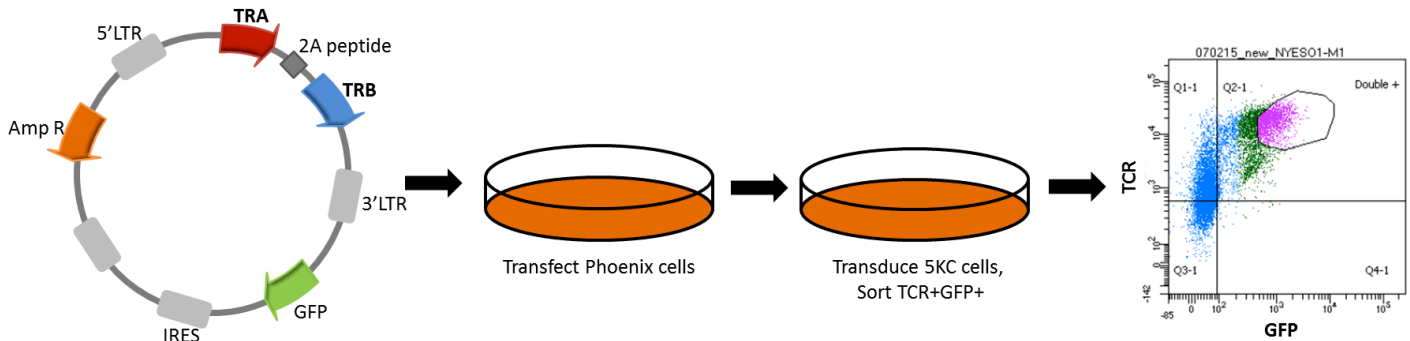


Figure 5. Production of TCR Transfectomas. Selected TRAV and TRBV are cloned into a retrovirus that contains the TRAC and TRBC regions. Retroviral vector is used to transfect phoenix cells using calcium phosphate method to generate virion containing medium. 5KC cells, a T cell hybridoma that lack TCR expression and express human CD8, are transduced with the newly produced virus. 5KC cells that show high expression of TCR and GFP are sorted by flow cytometry.

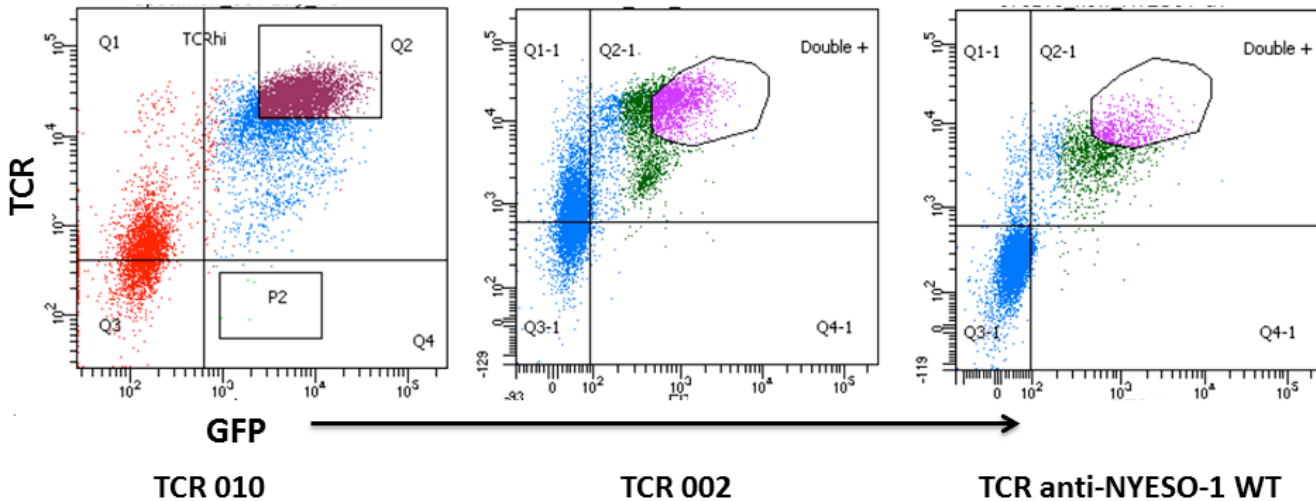


Figure 6. Selection of transfectomas with high TCR expression. 3 to 4 days after transduction, 5KC cells are stained with anti-TCR beta antibody and then evaluated by flow cytometry. Cells with higher expression of TCR and GFP are bulk sorted into tubes containing medium. Cells are cultured and expanded for several days before CD3 stimulation (as shown in **Figure 7**).

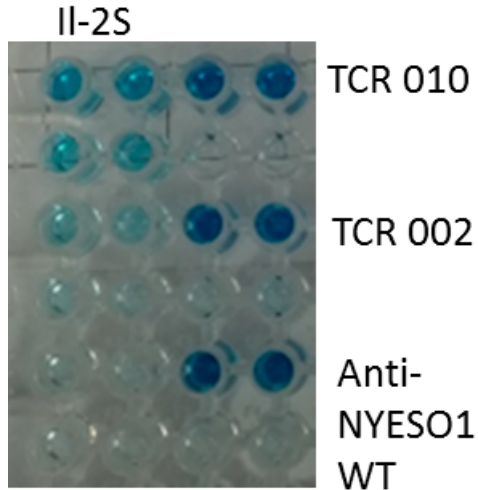


Figure 7. ELISA used to measure IL-2 production by 5KC transfectomas to CD3 stimulation. IL-2-specific antibody coated on a 96-well plate overnight. Standards and samples are added and incubated for two hours. Biotinylated anti-IL-2 antibody is added. After washing away unbound biotinylated antibody, HRP conjugated streptavidin is added. A TMB substrate solution is used to develop the color and the wavelength is detected at 650 nm. The wells are washed 3 to 5 times among steps.

Table 3. Summary of progress in the process of production of transfectomas.

Shared TCRs Selected	TCR 001	TCR 002	TCR 007	TCR 010	Anti-NYESO1 WT	Anti-NYESO1 HA	Anti-Her2
Cloning of V alpha and V beta into Retroviral vector	X	X	X	X	X	X	X
Transfection of phoenix cells		X		X	X		
Transduction of 5KC cells		X		X	X		
Select and sort cells by TCR expression		X		X	X		
Test for IL-2 production to CD3 stimulation		X		X	X		
Test for IL-2 production to APC + peptide							
Test for IL-2 production to BC cell lines							

Task 6. Generate MHC/peptide baculovirus libraries: months 1-40, Slansky

6a. HLA-A*02:01

We constructed an HLA-A*02:01 peptide library following the methods outlined previously. The peptide library sequence constraints are X(VTPLIA)XXXXXX(LV), yielding a potential library size of 1.54×10^{10} possible peptide combinations. PCR amplicons were generated from the baculoviral DNA which encodes the library and sequenced on the HiSeq2500. **Figure 8** shows the distribution of amino acids at each position as well as the expected amino acids at each position following analysis of ~330 million sequence reads encoding 11 million peptides.

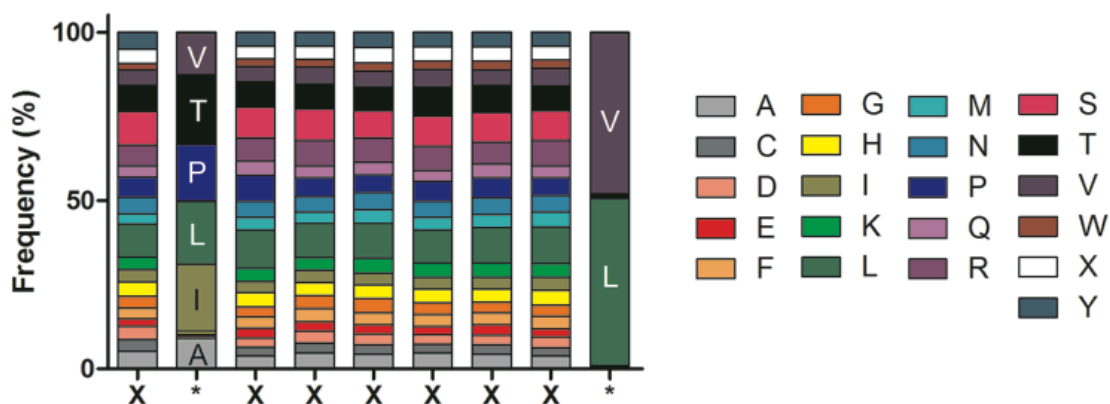


Figure 8. Sequencing of PCR amplicons of the peptide coding region of purified baculovirus DNA was performed to determine the amino acid distribution along the peptide. ~330 million DNA sequences were trimmed to the peptide sequence and translated. Unique peptide sequences were piled and the frequency of a given amino acid at position determined. Each amino acid is color-coded and although not visible every time, all 20 amino acids (as well as stop codons, X) are represented at the randomized positions. At the set positions, small numbers of contaminating amino acids were found due to mutation either from sequencing or from the viral polymerase during viral replication.

To screen these libraries, we must have soluble versions of the TCRs of interest. Toward this goal, we are generating the soluble TCRs from Table 4. **Figure 9** describes how the soluble TCRs are made and **Figure 10** shows that we have successfully produced TCRs 001, 002, and 010. **Table 4** summarizes the progress we have made generating soluble TCRs.

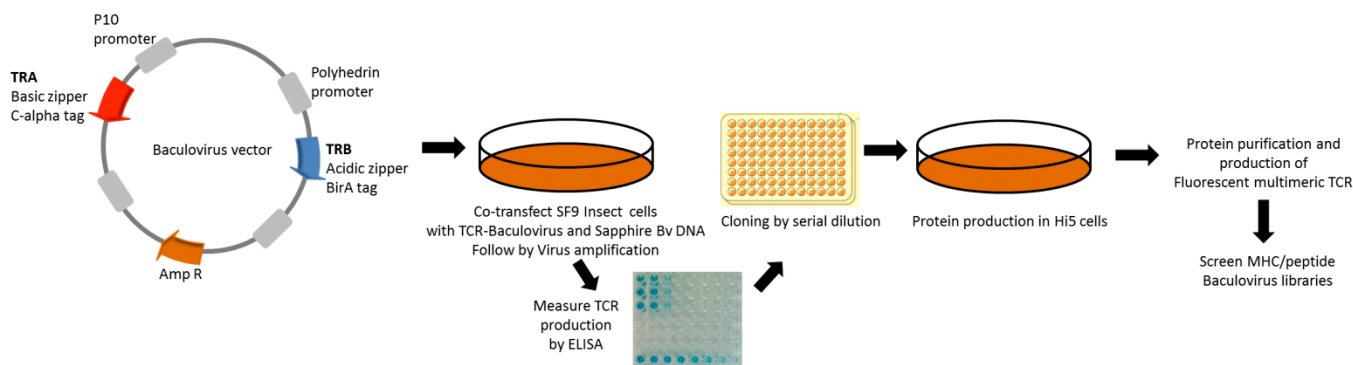


Figure 9. Production of soluble TCRs. Selected shared TRAV and TRBV from Figure 4 are cloned into a baculovirus vector that contains the mouse TRAC and TRBC regions. Baculovirus vector is used to co-transfect SF9 insect cells. Co-transfection of SF9 cells is initially assessed by survival of the cells compared to un-infected controls. The viruses are amplified using SF9 cells for 7 to 10 days and then TCR production is measured by an ELISA assay. Viruses are cloned by serial dilution. Clones with higher TCR production are amplified again and used to infect Hi5 insect cells for protein production.

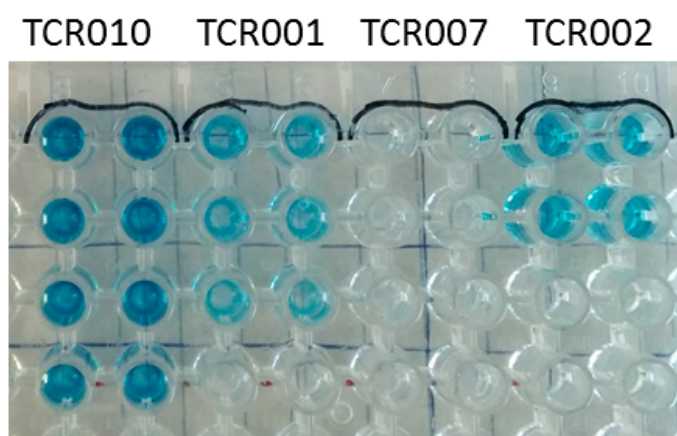


Figure 10. ELISA used to measure the amount of TCR produced by baculovirus infected insect cells. TRAC specific antibody coats a 96-well plate overnight. Standards and samples are added into the wells and incubated for two hours. Biotinylated anti-TRBC antibody followed by HRP conjugated streptavidin and TMB substrate solution are used as described for the IL-2 ELISA assay.

Table 4. Summary of progress in the process of production of soluble TCRs.

Shared TCRs Selected	TCR 001	TCR 002	TCR 007	TCR 010	Anti-NYESO1 WT	Anti-NYESO1 HA	Anti-Her2
Cloning of V alpha and V beta into Baculovirus vector	X	X	X	X	X	X	X
Co-transfection of SF9 insect cells	X	X	X	X	X		
Amplification of Virus	X	X		X			
Production of soluble TCR	X	X		X			
Cloning by limiting dilution	X	X		X			
Infection of Hi5 insect cells	X			X			
Protein purification							
Production of fluorescent multimeric TCR							

4. KEY RESEARCH ACCOMPLISHMENTS

- We obtained patient samples from the University of Colorado and Lee's Team at the City of Hope and used them to identify a panel of shared TCRs.
- We developed a rapid high-throughput low-cost single-cell emulsion RT-PCR method for determination of alpha-beta TCR pairs.
- We worked with the Spellman team at OHSU to develop an algorithm that rejoins the sequences obtained from the same cell. This program is called CompleteTCR and will be publically available.
- We showed that some TCR sequences in the tumor are enriched versus those found in the blood or lymph nodes.
- Using the emulsion PCR method, we fully analyzed the TCR sequences from CD8+ TILs and PBMCs from 20 HLA-A2+ bc patients and PBMCs from 7 controls.
- Analysis of these sequences showed that multiple TCRs are shared among patients.
- We generated a new HLA-A2-peptide library for screening. This library is randomized at all positions other than p2 and p9, the MHC-anchor residues. A sample of the library was sequenced, of 330 million different sequences/11 million peptides were identified. Although this does not show full saturation of the library, all amino acids can be detected in the randomized positions.
- We generated TCR transfectomas and soluble TCRs to screen various antigen libraries to determine their cognate antigens.

5. CONCLUSION

In summary, in the last year, we have finished developing a rapid, high throughput, low-cost, single-cell emulsion RT-PCR method to identify dominant clonotypes that are found in the tumor but not to the same extent in matched PBMC samples. Soluble versions of these dominant TCRs will be used to screen (1) known antigens and if these antigens are all negative, (2) a baculovirus-encoded peptide/MHC library. The antigens that the TCRs identify can be added to the pool of "known antigen" to reduce the number of TCR clonotypes necessary to screen the library with. In the next year with a one year no-cost extension, the remaining part of this epitope/mimotope discovery procedure will be optimized to screen and boost the T cell responses detected in bc samples.

6. PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS

PUBLICATIONS

Nothing to report, however, manuscripts describing (1) the TCR sequencing of alpha beta TCRs (separately) from bc patients and (2) the emulsion PCR method described in the first half of this report are in progress.

PRESENTATIONS

Nov 8, 2014, Analysis of the T cell repertoire in breast cancer using emulsion single cell RT-PCR, Talk in Barriers to Successful Vaccines session at the annual meeting for the Society for Immunotherapy of Cancer (SITC) National Harbor, MD (invited by Jonathan Bramson).

Nov 11, 2014, Elimination of the bottlenecks in T cell receptor antigen discovery, Immunology Forum Seminar, Johns Hopkins University, Baltimore MD (invited by Jonathan Schneck).

Oct 23, 2015 Analysis of the T cell receptor repertoire in breast cancer, seminar in the Integrative Biology Department, CU Denver, CO (invited by Raibatak Das).

7. INVENTIONS, PATENTS AND LICENSES

Nothing to report.

8. REPORTABLE OUTCOMES

- A. The most significant scientific advance and research tool that makes a meaningful contribution toward the understanding of how T cells of immune system interact with breast cancer was the development of the emulsion PCR method. With this method we can amplify the alpha and beta chains of the TCR together, which is necessary for the future experiment underway to identify the antigens that are recognized by these T cells.
- B. Using the emulsion PCR method, we identified a panel of TCRs that are enriched in breast cancer tumors (details in Figure 4).

9. OTHER ACHIEVEMENTS

Nothing to report.

10. REFERENCES

1. Ali, H. R., Provenzano, E., Dawson, S. J., Blows, F. M., Liu, B., Shah, M., Earl, H. M., Poole, C. J., Hiller, L., Dunn, J. A., Bowden, S. J., Twelves, C., Bartlett, J. M., Mahmoud, S. M., Rakha, E., Ellis, I. O., Liu, S., Gao, D., Nielsen, T. O., Pharoah, P. D., and Caldas, C. (2014) Association between CD8+ T-cell infiltration and breast cancer survival in 12,439 patients. *Ann Oncol* **25**, 1536-1543
2. Lee, A. H., Gillett, C. E., Ryder, K., Fentiman, I. S., Miles, D. W., and Millis, R. R. (2006) Different patterns of inflammation and prognosis in invasive carcinoma of the breast. *Histopathology* **48**, 692-701
3. Linnemann, C., Mezzadra, R., and Schumacher, T. N. (2014) TCR repertoires of intratumoral T-cell subsets. *Immunol Rev* **257**, 72-82
4. Jang, M., Yew, P. Y., Hasegawa, K., Ikeda, Y., Fujiwara, K., Fleming, G. F., Nakamura, Y., and Park, J. H. (2015) Characterization of T cell repertoire of blood, tumor, and ascites in ovarian cancer patients using next generation sequencing. *Oncoimmunology* **4**, e1030561
5. Buhrman, J. D., Jordan, K. R., Munson, D. J., Moore, B. L., Kappler, J. W., and Slansky, J. E. (2013) Improving antigenic peptide vaccines for cancer immunotherapy using a dominant tumor-specific T cell receptor. *J Biol Chem* **288**, 33213-33225
6. Howie, B., Sherwood, A. M., Berkebile, A. D., Berka, J., Emerson, R. O., Williamson, D. W., Kirsch, I., Vignali, M., Rieder, M. J., Carlson, C. S., and Robins, H. S. (2015) High-throughput pairing of T cell receptor alpha and beta sequences. *Science translational medicine* **7**, 301ra131

7. Turchaninova, M. A., Britanova, O. V., Bolotin, D. A., Shugay, M., Putintseva, E. V., Staroverov, D. B., Sharonov, G., Shcherbo, D., Zvyagin, I. V., Mamedov, I. Z., Linnemann, C., Schumacher, T. N., and Chudakov, D. M. (2013) Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol* **43**, 2507-2515
8. Bolotin, D. A., Shugay, M., Mamedov, I. Z., Putintseva, E. V., Turchaninova, M. A., Zvyagin, I. V., Britanova, O. V., and Chudakov, D. M. (2013) MiTCR: software for T-cell receptor sequencing data analysis. *Nature methods* **10**, 813-814
9. Warren, R. L., Freeman, J. D., Zeng, T., Choe, G., Munro, S., Moore, R., Webb, J. R., and Holt, R. A. (2011) Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome research* **21**, 790-797
10. Robins, H. S., Srivastava, S. K., Campregher, P. V., Turtle, C. J., Andriesen, J., Riddell, S. R., Carlson, C. S., and Warren, E. H. (2010) Overlap and effective size of the human CD8⁺ T cell receptor repertoire. *Science translational medicine* **2**, 47ra64
11. Sidney, J., Peters, B., Frahm, N., Brander, C., and Sette, A. (2008) HLA class I supertypes: a revised and updated classification. *BMC immunology* **9**, 1